DOCUMENT RESUME

ED 144 020 CS 003 669

AUTHOR Woods, William A.
TITLE Multiple Theory Formation in High-Level Perception. Technical Report No. 38.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Apr 77
CONTRACT 400-76-0116
NOTE 58p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Artificial Intelligence; Artificial Speech; *Cognitive Processes; *Models; Readability; *Reading Processes; *Reading Research
IDENTIFIERS Center for the Study of Reading (Illinois); *Inference

ABSTRACT

This paper is concerned with the process of human reading as a high-level perceptual task. Drawing on insights from artificial-intelligence research--specifically, research in natural language processing and continuous speech understanding--the paper attempts to present a fairly concrete picture of the kinds of hypothesis formation and inference processes that must take place during the reading process. It makes the case that many more alternative hypotheses are formed and evaluated than one might expect and that the majority of such processing happens below the level of conscious introspection. Two text passages of different levels of difficulty are considered in detail, and the applicability of such artifical-intelligence insights to the modeling of human reading are discussed. (Author)

CENTER FOR THE STUDY OF READING

Technical Report No. 38

MULTIPLE THEORY FORMATION IN HIGH-LEVEL PERCEPTION

William A. Woods

Bolt Beranek and Newman Inc.

April 1977

University of Illinois
at Urbana-Champaign
1005 West Nevada Street
Urbana, Illinois 61801

BBN Report No. 3580

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

The research reported herein was supported in part by the National Institute of Education under Contract No. MS-NIE-C-400-76-0116. A version of this paper will appear in R. J. Spiro, B. C. Bruce, and W. F. Brewer (Eds.), Theoretical Issues in Reading Comprehension, Hillsdale, N.J.: Lawrence Erlbaum Associates.

Multiple Theory Formation in High-Level Perception

W. A. Woods
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Ma. 02138

Abstract

This paper is concerned with the process of human reading as a high-level perceptual task. Drawing on insights from Artificial Intelligence research -- specifically research in natural language processing and continuous speech understanding -- the paper attempts to present a fairly concrete picture of the kinds of hypothesis formation and inference processes that must take place during the reading process. It makes the case that many more alternative hypotheses are formed and evaluated than one might expect and that the majority of such processing happens below the level of conscious introspection. Two text passages of different levels of difficulty are considered in detail, and the applicability of such Artificial Intelligence insights to the modeling of human reading are discussed.

1. Overview

High-level perceptual tasks such as reading, speech understanding, and visual scene interpretation are characterized by the need to discover a structured interpretation that accounts for the stimuli present. This process is prerequisite to deciding what has been perceived, and thus precedes whatever process decides what to do with the resulting perception -- what significance to attach to it, whether to remember it, how to incorporate it into the knowledge base of the perceiver, etc. In this paper, I will attempt to make the case that the process of arriving at an interpretation of the input involves the formation and evaluation of many alternative partial hypotheses about the input and that this process goes on largely below the level of introspective awareness of the perceiver. Even though skilled reading involves a variety of "metacognitive" strategies [Brown, 1978], in normal reading these processes are themselves invoked without conscious attention to the process of doing so.

I will focus on the problem of reading, and will draw on insights and analogies from work in Natural Language Parsing and Continuous Speech Understanding. Since I do not have space here to give an adequate introduction to all of the background material that I would like to use, I will instead refer the reader to three previous papers: "Meaning and Machines" [Woods, 1973b], "Syntax, Semantics, and Speech" [Woods, 1975a] and "What's in a Link: Foundations for Semantic Networks" [Woods, 1975b]. The most recent material on speech understanding is unfortunately only contained in technical reports. I recommend in particular the BBN final report [Woods et al., 1976], Vols. I, III, IV, and V. For a brief overview of the BBN speech understanding system, see Wolf and Woods [1977].

2. Multiple Hypotheses, Backtracking, and Nondeterminism

Many of our intuitions about techniques for dealing with multiple hypotheses come from work in parsing algorithms for formal grammars. This is an area where both theoretical results in automata theory and empirical results from programmed parsers are available. Other areas where such insights can be gained are formal theorem proving, machine vision, and continuous speech understanding. Perhaps the earliest and most widely known mechanism for handling multiple hypotheses is the predictive analyzer, or pushdown store parser [Kuno & Oettinger, 1963], using backtracking to handle the possibility of multiple alternatives.

A pushdown store parser is essentially an algorithm for analyzing a sentence in terms of rules such as "a sentence consists of a noun phrase followed by a verb phrase". It operates by using a stack or "pushdown store" of words and phrases that are expected to occur in the sentence and at each step compares the next input word against the top item on the stack, either finding a match between the current word and the predicted category, finding that the current word is incompatible with the predicted category, or expanding a predicted phrase type to a sequence of words and phrases that would realize it (in which case the new predicted categories replace the one from which they are derived on the stack). The stack operates very much like the spring loaded stacks of dishes and trays that one ~~finds in~~ some cafeterias (from which it derives its name). Backtracking refers to the process of saving enough information before making a choice among alternatives (in this case before choosing a particular rule to expand a phrase type) so that at a later time the situation prior to the choice can be reconstructed and a different choice selected.

For example, if a backtracking parser encounters a word that can be either a noun or a verb in a context where both nouns and verbs would be acceptable (albeit with different interpretations) then enough information is saved to remember the current position in the sentence and all of the decisions made so far before making one of the choices, say treating the word as a noun. Subsequently if this choice does not lead to an acceptable complete parsing (or in any case if all possible parsings are required), the saved information is restored and the other choice is tried.

Backtracking is not the only way of implementing an algorithm for exploring multiple alternatives, however. Other methods involve the creation of separate virtual processes for each of the alternative choices, which can then be run in parallel or in some prioritized order. (A virtual process behaves conceptually as if it were an ongoing process, but may in fact be competing for time and resources with other such processes so that the number of processes actually receiving resources at any given moment may be fewer than the number of "virtual" processes that are conceptually "active".)

A general method for specifying and discussing algorithms for this kind of processing of multiple alternatives is to treat the algorithms as procedures one of whose basic instruction types has the meaning "choose one of the following alternatives: ...". Such procedures are "executed" by an interpreter that systematically considers all possible combinations of such choices, either by backtracking or by the method of separate virtual processes. Such algorithms are referred to as nondeterministic algorithms, not because their eventual behavior is not determined, but because the specification of the exact order in which the various alternatives will be considered and the mechanism by which the combinations of choices will be enumerated are not overtly specified as part of the algorithm.

Parsing natural language is fundamentally a nondeterministic process. That is, it is not possible to make all of the correct decisions based solely on local information. Rather, tentative decisions must be made and explored to see if their consequences are plausible. For example, the decision as to whether a given word is being used as a noun or a verb depends on whether a complete parsing can be found under one or the other of the two assumptions. Thus "Fly" in "Fly fishing is fun" is a noun (used as a modifier), while in "Fly fishing equipment to Lakeview Lodge as soon as possible" it is a verb. (If you required a few seconds to find the correct interpretation of this sentence due to an initial incorrect interpretation of "Fly fishing," then you have experienced what is known as the "garden path" phenomenon.) Occasionally, several choices will lead to a complete parsing, in which case the sentence is ambiguous (as in "The girl guides fish", which could either be a statement about what girl guides do in their spare time or a description of a shepherdess for fish).

Predictive parsing is an intuitively satisfying analogy for thinking about human sentence processing since it seems to correspond well to our introspective awareness of stages of the parsing process. Notably, if the sentence is stopped at some point we can generate a completion; in general, we know what kinds of words to expect next at each point; and occasionally we are aware of having made a false early decision and having to back up to look for another alternative (in garden path sentences like the one above). One soon gets into difficulty with this analogy, however, since systematic predictive analysis for a natural language grammar makes many more extraneous tentative excursions down false parsing paths than people seem to. Also, the time required by people to recover from a garden path seems to be much shorter than that required by a backtracking parser [Collins & Quillian, 1971]. In the Artificial Intelligence field, this has given rise to a search for techniques to make the enumeration of parse paths more selective (e.g., by using "semantic" information) in order to correspond more closely to the level of backtracking that we are aware of when we read. In other quarters, it has given rise to the labeling of such processes as "Gestalt", treating the process a black box whose internal workings one does not attempt to understand.

In this paper I will attempt to make the case for another account, namely that the amount of partial hypothesis formation that a person actually performs in such tasks is far more than it seems, and that the bulk of it occurs below the level of introspective awareness. I will draw largely on experience in constructing speech understanding systems. In the speech understanding task, it is relatively easy to make the case that there is a great deal of hypothesis formation and testing required even to know what words have been heard, much less how they are to be parsed and interpreted. However, most of this processing takes place exceedingly fast and without conscious effort. Once it is realized that the human perceptual processes are capable of making large numbers of alternative hypotheses and choosing among them on the basis of relatively high-level plausibility judgments - all without conscious effort or awareness - it is not difficult to imagine such capabilities being used in reading and other high-level perceptual tasks.

3. Characteristics of the Speech Understanding Process

A naive view of speech understanding might consider it as a process of successively recognizing speech sounds (called phonemes), grouping phonemes into words, parsing word sequences into sentences, and finally interpreting the meanings of those sentences. However, considerable experience now indicates that the acoustic evidence present in the original speech signal is not sufficient to support such a process [Woods and Makhoul, 1974]. For sentences recorded from continuous speech, it is not generally possible to reliably determine the phonetic identity of the individual phonemes (or even to be sure how many phonemes are present) using the acoustic evidence alone. Experiments in spectrogram reading [Klatt and Stevens, 1971] indicate that the reliability of such determinations can be increased by use of the redundancy provided by knowledge of the vocabulary, the syntax of the language, and semantic and pragmatic considerations.

Tape splicing experiments [Wanner, 1973] seem to indicate that this low-level acoustic ambiguity is an inherent characteristic of speech and not just a limitation of human spectrogram-reading. Specifically, intelligibility of individual words excised from continuous speech is very low, but the intelligibility increases when sequences of two or three words are excised. It appears that the additional constraint of having to make sense in a larger context begins to resolve the ambiguities that were present when only the acoustic evidence was considered. This processing, however, happens below the level of introspection and has all of the subjective characteristics of a wholistic or Gestalt phenomenon. That is, if a sufficiently long sequence of continuous speech is heard, its correct interpretation usually appears immediately and effortlessly, without conscious awareness of the details of the process. The vast majority of our spoken communications are understood in this manner, and it is markedly contrasted with those cases where an utterance is garbled sufficiently to invoke conscious effort to decide what was said.

4. Theories, Monitors, Notices, and Events - A Computational Framework for Perception

The BBN speech understanding system [Woods et al., 1976; Wolf and Woods, 1977] has evolved within a general framework for viewing perceptual processes. Central to this framework is an entity called a theory. A theory represents a particular hypothesis about some or all of the sensory stimuli that are present. Perception is viewed as the process of forming a believable coherent theory which can account for all the stimuli. This is arrived at by successive refinement and extension of partial theories until a best complete theory is found.

In general, a high-level perception process requires the ability to recognize any member of a potentially infinite class of perceptible objects that are constructed out of elementary constituents according to known rules. That is, the object perceived is generally a compound object, constructed from members of a finite set of elementary constituents according to some kind of well-formedness rules. These elementary constituents, as well as the relationships among them that are invoked in the well-formedness rules, must be directly perceptible. Thus, a perceptual system must incorporate some basic epistemological assumptions about the kinds of things that it can perceive and the rules governing their assembly. The well-formedness rules can be used to reject impossible interpretations of the input stimuli, and may also be useable to predict other constituents that could be present if a given partial theory is correct.

This perception framework assumes mechanisms for using subsets of the input stimuli to form initial "seed" hypotheses for certain elementary constituents (stimulus-driven hypothesization) and mechanisms for deriving hypotheses for additional compatible constituents from a partial theory (theory-driven, or predicted, hypothesization*). It also assumes mechanisms for verifying a hypothesis against the input stimuli and evaluating the well-formedness of a compound hypothesis to assign it some measure of quality and/or likelihood. A theory may therefore be thought of as a hypothesis that has been evaluated in this way and assigned a measure of confidence.*

In the case of speech understanding, a theory can range from an elementary hypothesis that a particular word is present at a particular point in the input (a word match) to a complete hypothesis of a covering sequence of words with a complete syntactic and semantic interpretation. (In general, a theory can be a set of compatible word hypotheses with gaps between them and with partial syntactic and semantic interpretations.) A partial theory may be able to generate predictions for appropriate words or classes of words either adjacent to the words already hypothesized, or possibly elsewhere in the utterance.

Predictions are dealt with in our computational framework by two kinds of devices: monitors, which are passively waiting for expected constituents, and proposals, which are elementary hypotheses that are to be evaluated against the input. Proposals result in actively seeking stimuli that would verify them, while monitors passively wait for such hypotheses to be formed. The functioning of monitors assumes that there is an organizing

---

*Our notion of stimulus-driven hypothesization is essentially the same as that of "bottom-up" processing referred to in many discussions of such processes. However, our notion of theory-driven hypothesization is slightly different from the sense usually given to "top-down" processing in that it does not necessarily imply any global ("topmost") hypothesis, but only predictability by some other hypothesis, which may itself have been derived "bottom-up". The terms "top-down" and "bottom-up" in this sense come from the literature on formal parsing algorithms.

structure into which all derived partial hypotheses are placed as they are discovered and that the monitors can essentially set "traps" in this structure for the kinds of events that they are watching for. This is to be contrasted with continuous parallel evaluation of special processes (frequently called "demons") to watch for expected patterns in the input stream. Monitors perform no computation until and unless some other process makes an entry of the kind they are waiting for in some data structure.

The functioning of monitors is illustrated by an early speech understanding system dealing with concentrations of chemical elements in lunar rocks. There, for example, a word match for "concentration" would set monitors on the concept nodes for SAMPLE and CHEMICAL ELEMENT in the semantic network. If a word such as "Helium" was subsequently found anywhere else in the utterance, a check in the semantic network starting with Helium would lead to the superset category CHEMICAL ELEMENT where it would wake up the monitor from "concentration", thus detecting the coincidence of a detected hypothesis and a predicted hypothesis [Nash-Webber, 1975].

When a monitor is triggered, an <u>event</u> is created calling for the evaluation of a new hypothesis and the creation of a new theory if the hypothesis is not rejected. In general, a number of events are competing for service by the processor at any moment. In human perception, there may be full parallel processing of such events, but in a serial machine, these events must be queued and given processing resources on the basis of some priority ordering. (Even in human perception, there is probably some sort of priority allocation of resources, since various kinds of interference can occur.) In our computational framework, events are maintained on a queue in order of priority, the top event being processed at each step.

The processing of an event can result in new proposals being made, new monitors being set, and existing monitors being triggered to produce new events. Since so much hinges on the event chosen for processing, a major issue is that of assigning priorities to events in order to find the most likely interpretation of the input. In the BBN system, priority scores are assigned on the basis of Bayesian estimates of the probabilities of the competing theories, and certain control strategies and priority scoring metrics can be guaranteed to discover the most probable interpretation of the input.

5. Control Strategies

The above discussion leaves open issues such as when should seeds be formed, how many should be considered, should all seeds be worked on in parallel, etc. These issues we refer to as control issues. They have been critically important in computerized speech understanding systems. In the BBN system, for example, there are a variety of different control strategies that all fit within the above paradigm. Figure 1 illustrates one

#	SCORE	REGION	THEORY
1	0	11-14	ADD
2	0	4-7	NEED
3	-.455	0-3	SHOW !
4	-.605	12-17	TRIP
5	-.727	1-5	ROME
6	-.769	8-11	THERE
7	-1.25	12-18	TRIP-S !
8	-1.47	0-5	SHELLY
9	-1.65	15-17	END
10	-1.72	11-14	AND
11	-1.73	1-5	ANN
12	-1.74	0-5	CHEYENNE
13	-2.19	8-14	BERT
14	-2.26	2-6	ANY
15	-2.82	0-5	SOME

+15 ADDITIONAL EVENTS

Fig. 1. Seed events for middle-out strategy.

class of strategies in which seeds are formed anywhere in the utterance that sufficiently salient word matches are found. The figure shows the seed events formed as a result of an initial scan of an utterance for high likelihood word matches anywhere in the utterance. Each theory is assigned a score expressing its likelihood of being correct (actually a logarithm of the ratio of the likelihood of the acoustic evidence given the theory over the a priori likelihood of that evidence occurring independently). The region of the utterance covered by the theory is indicated by specifying its left and right boundary positions in a list of potential boundary positions (the left end of the utterance is numbered 0 and in this case the right end is numbered 18). The exclamation marks indicate the theories that are actually part of the correct interpretation.

For this general class of control strategies, referred to as "middle-out", theories are grown by starting with a seed word, asking a higher-level linguistic component to predict categories of words that can occur on either side of it, asking a lexical retrieval component to find the best matching words in those categories on the appropriate sides, and generating events for each such word found to extend the theory by adding that word. Thus, events will be placed on the event queue to add words both on the left and on the right ends of given theories. These "new word" events will compete with each other and with the remaining seed events on the basis of score to determine which event will be processed next, causing the processor to sometimes continue adding words to a given theory and at other times to shift its processing to a different competing theory.

Figure 2 shows the sequence of theories that are formed as a result of this process, starting with the event queue of Figure 1. (Brackets in the figure indicate theories that include the hypothesis that the left or right ends of the utterance have been reached. A number in parentheses after a theory number is the number of a preceding theory from which the indicated theory was formed by the addition of a new word.) Notice that the final theory is developed in this case by working independently on two different portions of the utterance starting from the seeds "show" and "trips". The final theory in figure 2 is in fact derived from a kind of event called a <u>collision</u> event which combines the theories "show me" and "trips" when they both notice the word "her" filling the gap between them. This event is formed during the processing of theory 13, although its score is such that it does not reach the top of the queue until theory 23.

Figure 3 shows the seed theories for a hybrid strategy in which seeds are started within a bounded distance from the left end of the utterance, and are grown right-to-left until they reach the left end, after which the remainder of the processing is left-to-right. Figure 4 shows the sequence of theories developed in the course of understanding this utterance using the hybrid strategy. The basically left-to-right nature of the

THEORY#1	ADD
THEORY#2	NEED
THEORY#3	SHOW !
THEORY#4(3)	[SHOW !
THEORY#5(4)	[SHOW ALL
THEORY#6(3)	SHOW ALL
THEORY#7	TRIP
THEORY#8	ROME
THEORY#9(4)	[SHOW ME !
THEORY#10(3)	SHOW ME !
THEORY#11(7)	HER TRIP
THEORY#12	TRIP-S !
THEORY#13(12)	TRIP-S] !
THEORY#14(13)	HER TRIP-S] !
THEORY#15(12)	HER TRIP-S !
THEORY#16	SHELLY
THEORY#17(16)	[SHELLY
THEORY#18	[SHOW ME HER TRIP
THEORY#19	SHOW ME HER TRIP
THEORY#20(11)	HER TRIP IS
THEORY#21(20)	HER TRIP IS]
THEORY#22(20)	OF HER TRIP IS
THEORY#23(9,13)	[SHOW ME HER TRIP-S] !

Fig. 2. Theories formed for middle-out strategy.

#	SCORE	REGION	THEORY
1	3.53	1-5	WHO
2	1.92	3-6	WE !
3	0.0	0-1	-PAUSE- !
4	-2.43	2-3	A
5	-3.24	5-10	ELEVEN
6	-4.32	5-9	IRAQ
7	-5.36	1-3	HER
8	-6.00	1-4	WHOLE
9	-6.18	1-5	DO !
10	-6.21	3-6	WERE
11	-6.53	3-7	WORK
12	-6.85	1-4	HIS
13	-7.00	1-5	HOW
14	-7.12	1-6	HAWAII
15	-7.21	3-6	WHERE

+39 ADDITIONAL EVENTS

Fig. 3. Seed events for hybrid strategy.

```
THEORY#1          WHO
THEORY#2(1)       [WHO
THEORY#3           WE !
THEORY#4(3)       DO WE !
THEORY#5(4)       [DO WE !
THEORY#6(3)         [WE !
THEORY#7(5)       [DO WE HAVE !
THEORY#8(7)       [DO WE HAVE A !
THEORY#9          -PAUSE- !
THEORY#10(9)      [-PAUSE- !
THEORY#11(3)        ARE WE
THEORY#12(8)      [DO WE HAVE A SURPLUS !
THEORY#13(12)     [DO WE HAVE A SURPLUS] !
```

Fig. 4. Theories formed for hybrid strategy.

hybrid strategy, except for a bounded initial delay in getting started, seems to be a reasonable possibility for a model of human speech understanding, since it is clear that human processing of speech does not involve the buffering of a complete sentence before understanding begins. The reading process, however, may involve more jumping around in a text and share some of the characteristics of the middle-out strategy.

6. Subconscious Processing of Alternatives in Speech and Reading

Essential in all of the above strategies is that at any given time there are a number of incomplete, competing possible interpretations, requiring a strategy to determine when processing resources should shift from one partial theory to another. One might initially suspect that a listener would be consciously aware of such competing possible theories and that shifting from one to another would correspond to the noticeable phenomena experienced with garden path sentences. However, our experience with speech understanding systems indicates that the construction and evaluation of competing partial hypotheses is far more prevalent than our introspective awareness makes apparent. It seems, then, that there must be some process for handling multiple alternative hypotheses that is subconscious and highly efficient (for perhaps a limited class of phenomena) and there is some other process that makes alternatives visible to our perception (perhaps for a more difficult class or phenomena or for those that have not been frequent enough to have been "compiled" into our subconscious process).

This distinction between conscious and subconscious processing is of course not an original observation. For example, Becker [1972] has referred to the former, subconscious processes as intermediate level cognition. It is likely that the bulk of our intelligent processing consists of this kind of subconscious intermediate level processing, and that the "thinking" which we are aware of at the conscious level is merely the tip of the iceberg. Conscious processing seems to be largely sequential and relatively slow, while the subconscious processes must (by virtue of the information processing tasks that they accomplish) be either highly parallel or exceptionally fast (or both). Other examples of the kinds of subconscious processes to which I refer are the retrieval of an association given a stimulus and the recognition of a face.

I will argue that the reading process contains large components of this kind of subconscious processing. In general, a reader is aware only of the final interpretation which he places on a sentence in a text; he is not aware of all of the intermediate stages of the derivation of that interpretation, all of the local ambiguities that were resolved by later context, and the accessing of factual memory to evaluate plausibility of competing partial interpretations. Nevertheless, a close examination of the information processing required to arrive at the final interpretation indicates that large amounts of such processing must be going on. By way of example, consider the resolution of the antecedent of the pronoun "they" in the following pair of sentences adapted from Winograd [1971]:

a) The city council refused to grant the women a parade permit because they feared violence.

b) The city council refused to grant the women a parade permit because they advocated violence.

This resolution appears to happen effortlessly, but the criteria that are needed to make the selection indicate that a very sophisticated inferential process is acting here below the level of awareness. It is not sufficient to choose the preferred interpretations ("council feared violence" in the first case and "women advocated violence" in the second) on the basis of some simple strength of association, between "council" and "fear" or between "women" and "advocate". If anything, one's a priori expectations for such associations would go the other way. It would appear that on a priori grounds, women fearing violence would be at least as plausible as city councils fearing violence. It is only at the level where one begins to evaluate what would be plausible grounds for a city council to make a refusal that the preference emerges. This implies a process in which very high-level evaluations of alternative interpretations of the sentence are required before fairly low level ambiguities are resolved. It does not seem possible to even formulate the necessary question to resolve the ambiguity until both possible interpretations have been formulated and elaborated at least to the level of justification of grounds for refusal. It is clearly not the case that a process must be simple and easy just because it happens in the head very rapidly and without apparent effort. In subsequent sections we will consider many more examples of this kind of inference in reading.

7. Implications for Psychological Models of Reading

The above picture of the perception process as involving many competing partial hypotheses that are formed and evaluated below the level of introspection has significant implications for the design of experiments to investigate the reading process. For example, it is possible that a large percentage of the reaction time for understanding many sentences is due to the evaluation of competing alternatives and not due at all to aspects of the correct interpretation. Thus, two different sentences or texts having some identified difference in the structure of their correct interpretation(s) that one would like to investigate may also have differences in reaction time due to extraneous differences in the number or complexity of competing partial hypotheses that are not part of the final interpretation.

To make matters worse, the effect of such competition depends critically on whether the process for handling alternatives is a serial or parallel process. As a simple example, consider a parallel processing control strategy which at each step extends the n most likely theories in the event queue for some number n (perhaps large). Then up to n different theories may be pursued as alternatives without introducing reaction time delay, but alternatives in excess of n would introduce delays.

The speech understanding models also predict that a favored competitor may be found much sooner than an unfavored competitor, and that the reaction time for a given interpretation may be longer if the correct interpretation is delayed by a partial theory that looks better locally but does not extend as well.

In the human brain, the implementation of such capabilities may not correspond to that of a queue sorted by priority. Other "implementations" could involve a large number of active memory elements that interact among themselves in such a way that the highest priority member is selected for processing, or resources could be allocated to all pending events in proportion to their priorities, etc. Thus, timing predictions cannot be made directly from the performance of computerized systems running on serial machines, but have to be made from an extrapolation of such behavior to hypothesized mechanisms in the brain. Unfortunately the number of degrees of freedom in such extrapolation is large, so working out a psychologically verifiable model of the process is likely to be lengthy with many intermediate models that have to be formulated and then rejected.

8. A Close Look at Some Reading Material

Having now given, I hope, sufficient reason to believe that both in speech understanding and in reading there are significant inferential processes that occur below the level of introspection, and that the characteristics of these processes are quite different from those of conscious inferential processes (especially with respect to apparent degree of perceived difficulty versus the amount of information processing actually done), I would like now to look in some detail at the reading process.

I will present fragments from two passages with significantly different levels of reading difficulty, accompanied by a detailed analysis of some of the inferential problems that must be solved for their full understanding. Each fragment comes from a passage of 4-6 paragraphs that is relatively self contained, but has apparently been extracted from some larger work. The passages are from the Riverside Research Institute Reading Competency Test [Riverside Research Institute, 1974].

8.a. Taffy

Taffy is a puppy. She is small. She is soft. She is sweet. What if Taffy belonged to you? Would you know what to do? Read this story. You will find out how to take care of Taffy.

Taffy sleeps a lot. She needs a place to rest. Take a large box and cut out one side. Put in something soft to lie on. This will be Taffy's bed. Put it in a warm place. Keep it dry and clean.

The apparent purpose of this story is to let a child know what having a puppy is like. This includes preparing the child to empathize with the puppy and take care of it. The story assumes knowledge of what a puppy is. It adds to the concept, if not already present, the attributes of smallness, softness, and sweetness. (Notice that the child must conclude that sweet is not a taste but a personality attribute.) Prominent in the story are rhetorical questions, designed to make the child think, to extrapolate from the given attributes. The story sets up the impression of holding a small, soft puppy. Then it asks rhetorically whether child would know what to do with a puppy. The intention is apparently to set the child up in the imaginary situation of having a puppy and not knowing what to do with it. The story then suggests a resolution of the dilemma -- namely, read the story. The result will be that child will know how to take care of a puppy.

Although the entire story is written in terms of a particular puppy, Taffy, it is clear that what the reader is to get from it is to apply to all puppies. How is this distinguished from, say, a story about Lassie, in which the story is intended to give attributes of a particular dog rather than dogs in general?

The rest of the story, of which the above is only a brief excerpt, is a conditional program for taking care of a puppy. It involves following directions. The story might be read for general information, in which case the goal of the story is for the child to remember at some later time something about taking care of puppies. It could also be read with an immediate need at hand and followed like a cookbook. The following is a sentence-by-sentence account of some of the characteristics of this story.

1. [Taffy is a puppy.]

   This sentence occurs with no prior context. In isolation, it would be an informing statement and would presume that the reader knew some referent for Taffy, but not that it was a puppy. As part of a story, perhaps it can be interpreted as introducing a character. Certainly, the intended result of the sentence in understanding the story is the creation of a temporary entity of type puppy, named Taffy, about which the reader expects to hear more in the course of the story.

2. [She is small.]

   This sentence adds an attribute to the introduced character (character development). The anaphoric reference ("she") here is easy, but in general would be more difficult if there were more than one potential antecedent. In this case, since Taffy is at best slightly correlated with feminine gender, the feminine pronoun is more likely to be interpreted as adding information to the model, rather than a restriction on the antecedent. This is an example of a situation where what would be a prerequisite in one context is adding information in another.

3,4. [She is soft. She is sweet.]

   Same as above.

5. [What if Taffy belonged to you?]

   This is a rhetorical question that exhorts the reader to add a representation of himself to the situation and draw some forward inferences. (ulterior motive -- arouse feelings)

6. [Would you know what to do?]

Another rhetorical question. This invites the reader to ask this question of himself. (ulterior motive -- reader concludes he wouldn't know what to do or at least would like some guidance or refreshment)

7. [Read this story.]

An imperative. Notice that this sentence has a self reference - the antecedent of "this story" is not present in the discourse, but refers to the story itself. The reader must be able to identify the reference. Note that a literal interpretation (e.g., by a computer) might be to begin again at the beginning to read the story, which would result ultimately in an infinite loop. The child is presumed to be sensible enough to know that he has already read part of it and can fulfill the command by merely reading the rest of the story. Alternatively, he can assume the phrase "this story" is a forward pointing reference to what remains, and that the part that has just been read is somehow a preamble or something, but not properly part of the story. The child could conceivably take either of these alternatives.

8. [You will find out how to take care of Taffy.]

A declarative sentence, but presumably interpreted as the consequent of a conditional -- "you will find out ... if you read the story". Not only that, but it says that you will learn how to take care of Taffy, when in fact, what you will learn how to do is take care of any puppy. In the entire story, Taffy is clearly the name of a variable that can later be instantiated. Does the child fully understand this? Does anything in the subsequent story depend on Taffy being female? How does the child determine at what level to generalize what he learns. E.g., to all puppies? to female puppies? to small, soft, female puppies? or perhaps to small, soft, animals? On the other hand, if he were really learning how to take care of Taffy, a particular dog, he would have to know where she lived, etc.

After the first paragraph, there is an expectation that a major portion of the rest of the story (maybe all of it) will be instructions in how to take care of a dog. We will discuss just a few more sentences.

9. [Taffy sleeps a lot.]

This adds more to properties of a puppy (Taffy?) (ulterior motive -- motivates next sentence and rest of paragraph).

10. [She needs a place to rest.]

An implicit "therefore" relates this sentence to the previous one. (ulterior motive -- creates a need which subsequent sentences will tell how to meet). Does a child pick up the implied causality connection between this and the previous sentence?

11. [Take a large box and cut out one side.]

This imperative is not necessarily to be done immediately as was the one in the first paragraph. If the reader is in recipe mode, he might do it immediately, but even then it would be better to incorporate the instruction into a plan that is being built up, and go on to read more instructions before doing anything. If in the mode of reading for future information, the reader should form the plan, and also elaborate the reasons for the individual steps, so that he will be more likely to remember it. The previous sentences motivating this sentence help in this respect, but the reader should also internally answer such questions as "Why a box?", "Why cut out one side?", and "What am I going to do with the box?" Also, he should understand that this use of "take" does not mean "steal". Likewise, he needs a decision on how large is large? This is one reason for incorporating the instruction into a tentative plan and then going on to elaborate the plan. Subsequent decisions may have a bearing on how large the box must be. For example, a refrigerator box or piano box is not necessary.

8.b. <u>Growth</u>

Growth is characterized by organization. Group growth charts show many age level uniformities and predictable age level changes. Differences between individuals in growth certainly exist. However, most individuals tend to be only slightly variable in their rate of growth from one age level to another.

This story is much more abstract and sophisticated than the Taffy story. It is apparently intended to impart to the reader some general information about the growth process for some indefinite future use. The passage could possibly be an excerpt from a biological text book. It is characterized by a much greater necessity for the reader to suspend judgment on possible interpretations of portions of the passage until he has read further. Among other things, it is characterized by a large number of noun-noun modification constructions ("group growth charts", "age level uniformities", etc.), where a noun is used to modify another noun. The intended meaning of such constructions is critically dependent on world knowledge of the habitual relationships between the two nouns. A detailed account of some of the difficulties of this passage follows:

1. [Growth is characterized by organization.]

This sentence introduces the topic "growth" and focuses on its organizational aspect. It doesn't say much as a factual assertion. If the reader asks himself a great deal about this question out of context, he may try to instantiate examples of kinds of growth and try to find instances of organization associated with them. In all likelihood, however, he will do little with this sentence, but go on to see what's coming next.

2. [Group growth charts show many age level uniformities and predictable age level changes.]

This is a difficult sentence syntactically. It has three noun-noun modifier constructions which require inference to determine their meaning. If the phrase "group growth charts" means something to the reader as a technical term in some field, then the sentence is considerably easier, but otherwise, he is left wondering "groups of what?" (possibly people, but the story doesn't say). The most general possible assumption is growth of organisms (or perhaps even more general still - including institutions, cultures, empires, etc. - all of this is consistent with the very abstract first statement, and it is possible that the vagueness and abstractness of the first sentence is intended to make such general interpretations of the second possible.)

The phrases "age level uniformities" and "age level changes" are again considerably more easy to understand if they are previously known technical phrases than if they have to be figured out from scratch. In the latter case, the reader needs to decide whether the phrases use the word "age" to modify the phrases "level uniformities" and "level changes" or whether the phrase "age level" modifies "uniformities" and "changes." In this case, we assume that the reader is familiar with the concept of "age level" and would for that reason choose the latter interpretation, since he is not likely to find a plausible reference for the phrase "level uniformities". (Note that "level changes" is a perfectly good phrase, but would require a subquestion -- "what kinds of levels"). The sentences would be more properly punctuated as "age-level uniformities" and "age-level changes", but such punctuation is frequently missing and a good reader is expected to be able to understand the phrases anyway. The difficulty is greater without the hyphens, since it opens up the possibility of different interpretations and requires the reader to resolve the ambiguity using semantic and pragmatic considerations.

3. [Differences between individuals in growth certainly exist.]

It's hard to say what this sentence does, except to set up expectations for what will come next. The choice of the word "individuals" suggests a restriction of the notion of growth being talked about to single organisms and perhaps to people, rather than the more sweeping generalizations to organizations and cultures that were possible for sentences 1 and 2. For a reader who jumped to the conclusion that growth of people was being talked about at the outset, this story may actually be easier than for the more intellectual reader who made the more general extrapolation and is now faced with the necessity to reconcile this use of the word "individuals" with his previous generalization.

It might be possible that this sentence just referred to certain kinds of growth (i.e., of individuals), while the story as a whole would still be about the more general concept of growth. However, some pragmatic rule seems to say that if that were the case, then the story would have contained some transition indicating a temporary shift to considering growth of individuals. Such a sentence might have been "In individuals, differences in growth exist" (with a presupposition that what is being said is not true for growth in other than individuals). In the absence of such a focus-shifting locution, the good reader should assume that this sentence is still about the same topic at the same level, and therefore his initial uncertainty about the generality of the concept of growth being discussed should here be resolved to refer to the growth of individuals (probably human).

Some processing load, which may have been considerable, was required to keep this decision open to this point, and some is still required until the reader determines whether the individuals of concern are people or individual organisms. The resource expenditure required to keep such decisions pending may be considerable, and may significantly affect reading diffioulty or the performance of resource limited individuals in reading such passages.

4. [However, most individuals tend to be only slightly variable in their rate of growth from one age level to another.]

The "however" here has the effect of negating or diminishing the strength of the preceding sentence. The general pragmatic rule seems to be that "however" will introduce a sentence (or paragraph) that contradicts some expectation that the reader might otherwise have drawn from the preceding sentence (or paragraph). It takes considerable inference to determine which specific expectation from the previous sentence is here being contradicted. One possibility is that the author is trying to contradict the inference that the differences referred to in the preceding sentence are large*. (This would explain the use of "certainly" in the preceding sentence, although I am not sure what the rule is, or whether the use of "certainly" leads the reader to expect something of this sort to follow).

The use of "most" implies that not all individuals satisfy the assertion. The use of "tend" is another hedge word that says the statement is not always true. The statement so qualified, that is, the predicate which tends to be true of most individuals, is that their growth rate is relatively stable from year to year. Thus, "however" is now explained not as contradicting the possible inference that the differences are large, but instead it contradicts the possible inference that differences in one dimension (from

*Note that I am not necessarily assuming that the reader consciously invokes models of the author and his intentions in understanding such sentences; the rules can be formulated in such a way that no author or ulterior motives need to be explicitly mentioned. However, it is probable that a reader who does have a specific model of authors and their intents is better able to understand such passages by invoking such a model, as discussed by Bruce [1978]. It is likely that the necessary rules can be more succinctly represented and learned by using the concept of an author and his intentions than by learning them by rote.

individual to individual) might indicate differences in another dimension (from year to year). To fully understand these two sentences and their uses of "certainly" and "however", the reader must have a concept of this kind of comparison.

Whether it is necessary for the reader to actually set up a model of something potentially varying over several dimensions in some space, and mapping age differences and individuation onto two such dimensions, I am not sure. However, it is clear that a reader must either ignore the cues given by the use of "certainly" and "however", in which case he risks misinterpreting some part of the story, or else he must have a sufficiently rich analysis of the sentences and their possible interpretations that he can fully account for them. A reader who is in between these extremes, who is trying to account for the use of the words, but is not adept at constructing possible underlying models to account for them, will find this passage more difficult than either of the other two kinds of readers. It would be a considerable challenge to devise a test that would distinguish these three different levels of reading. One complicating factor is that which of the three levels of competence a reader has mastered would probably vary considerably from one set of syntactic cues to another.

Notice that the purpose of this story seems to be to build up a fairly abstract model of how growth occurs. The reader is not asked to do anything immediately except somehow to assimilate the story. The purpose of the reader in reading the story thus becomes of paramount importance, and if his purpose is to prepare himself to perform in some way in the future using a model of how growth occurs, then he is likely to do different things in reading the story than if his purpose is, say, to be prepared to answer questions at the end of the passage, or to fill in missing blanks in the story. Of course, if this passage is read by an elementary school reader, it is likely that he does not have the goal to prepare himself for some specific future performance, since it is likely that he cannot even imagine a specific use for the information. However, the author may have such a use in mind in writing it. In particular, the author might be writing for an adolescent who may be concerned that his own rate of growth is too slow or too fast, and may want to convey such a growth model so that the reader can understand what's happening to him. It seems likely that two students, one of whom is concerned about his slow rate of growth, and another who is determined to become a biologist, will get completely different things out of this story.

9. Some Details of the Reading Process

The previous discussion has outlined some of the inferences that that need to be made in the course of understanding these two passages. However, it sheds little light on how those inferences might be organized and carried out. Experience with computerized language understanding programs and speech understanding systems can give us some insight into how these processes might happen, although the picture at the moment is far from complete. In the remainder of this paper I will discuss in more detail some of the low level decisions that have to be made and alternative hypotheses that have to be generated and considered, pointing out as I do so where existing techniques in natural language processing by computers have been developed to handle similar problems. I will consider in particular the first two sentences of the Growth passage.

The first sentence, "Growth is characterized by organization," is not difficult syntactically. It is a straightforward passive sentence with little potential for syntactic ambiguity. The first word can only be a noun, the second is unambiguously a verb (although it is not clear after only two words whether it is the main verb of a copular sentence or an auxiliary verb of a passive). The fact that the third word is a past participle resolves the sense of "is" to that of an auxiliary in a passive sentence. [Woods, 1970] gives a detailed account of how ATN grammars can be used to recognize and disambiguate this main-verb/auxiliary distinction. After this, "by" is unambiguously a preposition (although whether it is indicating the agent of the action or introducing something which an action takes place beside is not unambiguous). Organization is unambiguously a noun. Hence the discovery of the syntactic structure of this sentence is straightforward and involves little nondeterminism. Every local syntactic ambiguity is resolved by the immediately following word.

What the sentence means, however, is something else again. Both noun phrases in this sentence are mass nouns (nouns that can be used without determiners in the singular form as if they denoted a substance, as opposed to count nouns which can be counted). When used in the singular with no determiner, they are to be interpreted as refering to the general concepts that they name. ("A growth", "growths", or "an organization" would get completely different interpretations; the difference is flagged by the way the words are used syntactically.)

"Growth", as was pointed out previously, can name several different concepts that the reader might have in his head. If he has only one such concept, then the reading task is easier (although it may get difficult later on if the concept he has is not the one the author intended). If he has several, then the interpretation of the sentence is semantically ambiguous with respect to the reference of this phrase. This can be represented temporarily by associating with the noun phrase a list of alternative possible interpretations (such as is done in the semantic interpretation procedure of the LUNAR system [Woods, 1973a]).

The interpretation of "organization" appears to be somewhat different, largely because of its absence from surface subject position in the sentence. (Notice that "Organization characterizes growth" would not have the same effect as the initial sentence of this passage.) The correct interpretation of the sentence is an assertion about growth and what is being asserted is that it has many of the properties associated with organization. That is, "organization" names a concept where certain characteristics are to be found that are to be associated now with growth.

The differences in interpretation of the two noun phrases are due to their position in context as different arguments of the verb "characterize" and as different role fillers in the surface structure of the sentence. The LUNAR system handles such differences in interpretation as a function of context by providing context dependent parameters to the routine that computes possible interpretations of constituents. These parameters are used to determine the interpretation rules to be used for interpreting a constituent. (It is possible in general for different possible higher interpretations to call for the interpretation of a constituent in different ways, so provisions are required to keep alternative interpretations of different constituents coordinated with each other.)

For interpreting the clause as a whole, a semantic pattern for "characterize" (or perhaps for "is characterized by") is required. This is a fairly abstract and somewhat vague semantic relation. My edition of Webster's gives two senses of "characterize" -- (1) to describe the character or quality of: DELINEATE, and (2) to be a characteristic of: DISTINGUISH. Leaving aside the problems of the adequacy of such dictionary definitions, I interpret these definitions to focus on two senses of characterize -- one in which sufficient conditions are given as in "four equal length sides and right angles characterize a square", and the other in which merely prominent characteristics or only some characteristics are given. Which of these two senses is chosen would make a big difference in what the reader believed the passage to say. In the first case, it would say that anything that was organized would be growth - obviously false if one knows anything about growth and organization. The second interpretation is that organization is a prominent characteristic of growth. The reader must decide whether this sentence is trying to inform him of an astounding new fact or is merely asserting organization as a property of growth. Presumably the second choice is more plausible than the former.

In the LUNAR system such ambiguity of word sense was indicated by having several semantic interpretation rules associated with a given head word, both of which might match a given constituent being interpreted. Thus a procedure for generating both possible interpretations is straightforward. The problem of evaluating which of two interpretations is more plausible is more difficult, and no computer system at the moment makes such plausibility evaluations. Almost all current computer models of such processes make all-or-nothing decisions that an interpretation is either possible or impossible, with no shades in between.

One is tempted to say that the above choice is obvious (and implicitly therefore easy), but it is not clear exactly how far each of the two alternative interpretations has to be elaborated before the choice can be made. It seems to be necessary to actually formulate the erroneous interpretation and pose it as a question against one's knowledge in order to determine that it is false and therefore the other interpretation is to be preferred.

Consider here the understanding task imposed on a reader who was not aware of both senses of the verb "characterize". If he had only the correct interpretation, then the task would in fact be easier than for a more advanced reader who knew them both. On the other hand if he had only the wrong sense, then the first sentence would be apparently false, and he would have a difficult time with the passage. An additional possibility is that the reader doesn't really understand what "characterize" means and interprets this sentence merely as establishing some kind of association between growth and organization, which as it turns out, is about all that the sentence is intended to accomplish anyway. The primary role of this first sentence seems to be merely to establish growth as the topic, and perhaps bring certain aspects of growth into focus (the organizational aspects). Thus, the reader can get the appropriate effect without fully understanding the sentence at all.

The second sentence, "group growth charts show many age level uniformities and predictable age level changes", is far more complex syntactically as well as semantically. English syntax permits the use of nouns to modify other nouns in almost infinite profusion, but the interpretation of the meaning depends on non-syntactic world knowledge. Moreover when more than one such noun modifier is used, a structural syntactic ambiguity is introduced that requires world knowledge to resolve.

If "group growth" occurred in isolation, the structure of a noun, "group", modifying another noun, "growth", would be the only possible syntactic interpretation. The determination of what it means would depend on the ability of the reader to identify a plausible relation between the two words (in this case a group of things can grow, giving rise to an interpretation "growth of a group"). However, in "group growth charts" it is ambiguous whether "group" modifies "growth charts" or "group growth" modifies "charts".

The fact that charts are devices for depicting things gives "growth charts" the possible interpretation "charts depicting growth", and "group growth charts" the possible interpretation "charts depicting the growth of groups". However, the correct interpretation is probably "charts depicting growth by group", derived from "group" modifying "growth charts", requiring the reader to either know about or imagine a kind of growth chart that would be distinguishable by having something to do with groups. This meaning is not very distinct from one of the possible interpretations of "charts depicting growth of groups" (indicating that the meaning of that structure is far from unique), and there is certainly not enough evidence at this point in the passage to resolve which of these different syntactic structures or which of their semantic interpretations is to be taken as the author's intention. The possibility of this phrase having any of several possible interpretations needs to be held open until more of the sentence is processed.

The process of deciding that "charts" is the last word in the noun phrase and that "show" is the main verb requires some further local ambiguity, although the fact that there is no determiner on the noun phrase requires that it end either with a plural noun or a mass noun, so that "show" could only be included if it were followed by another noun. ("Show" could be a noun instead of a verb, although it would be difficult to put a plausible interpretation on "group growth charts show" as a noun phrase in this context.)

Exactly similar problems are encountered in interpreting "age level uniformities" and "predictable age level changes". The latter has even greater potential ambiguity due to the possibilities of "predictable" modifying "age", "level", or "changes" (e.g., "predictable changes in age level", "changes in the predictable age level", or "changes in the level of predictable age"). Somehow a reader makes a choice from among these different possible interpretations, usually without much conscious effort and usually correctly (or at least one of several equally acceptable interpretations). In this case, none of the above possibilities is correct, but instead, the thing that is doing the changing is elliptical (presumably growth or some growth parameter such as height or weight) and the correct interpretation is more like "predictable changes [in some growth parameter][as a function of] age level". Changes in a growth parameter as a function of age are certainly to be expected and therefore "predictable" whereas it is difficult to imagine anything on a growth chart that would correspond to a predictable age level (what would be doing the predicting). Hence, a probable role of "predictable" is to modify the concept "age level changes" as a whole, although probably as a non-restrictive modifier.

The above interpretation is not completely correct, since if a similar evaluation of "age level uniformities" is carried out, one is led to look for something that is uniform within an age level (again, presumably some growth parameter). The fact that conjunctions require some degree of parallelism between the two things being conjoined appears to demand that the role of the phrase "age level" should be the same in the two conjuncts. This slightly contradicts the otherwise well-motivated interpretation of "predictable age level changes" above. Instead, an interpretation in which something changes within an age level rather than as a function of age level is required to maintain this parallelism. This can be met by replacing the relation "as a function of age level" that was postulated above with "within an age level". This, however, removes the foundation for the argument justifying "predictable". Looking further for a different justification, one can suppose that the charts somehow show predictable changes (i.e., the charts do the prediction?), and that "predictable" is a restrictive adjective here telling something about the kinds of changes that the charts show.

Computerized parsing algorithms using formal grammars for extensive subsets of English can systematically enumerate all of the possible ways of grouping the words in such noun-noun modifiers, but the process is usually combinatorically expensive and few computer models deal with such constructions. Attempts to use semantic information to guide a parser to construct only possible interpretations have been attempted, but nothing that begins to match the complexity of the above discussion has currently been implemented. In general, the techniques for efficiently coupling syntactic and semantic knowledge in such situations are still being explored and the results are not in.

The discovery of the preferred interpretation in the above example requires the judgment that all of the various ways of grouping the individual words in this noun-noun modifier sequence have implausible interpretations in this context (what a group growth chart would show) without the addition of an additional elliptical participant in the underlying representation (namely what changes). The chain of reasoning justifying the correct choice is something like:

> "I know that a growth chart should show changes in some growth parameter (that's what growth is) and not changes in growth (at least not directly -- that's the derivative of what a growth chart would depict), so that must fill the "changee" role of the change being discussed. I recognize "age level" as a concept, so that is probably the role that "age" is filling and not a modifier of "level changes". ("level changes" is such a concept also, but I can find a plausible connection between "age level" and "changes" and cannot find one, or at least not a different one, between "age" and "level changes"). To relate "age level" to

> "changes [in a growth parameter]", I can take advantage of further knowledge (or imagination) about growth charts and speculate that the charts might be broken down by "age level". That could be the role that "age level" is filling here, but that would violate parallelism of conjunction with the previous phrase. I could either try to reinterpret the previous phrase to establish parallelism or I can try the same relation between "age level" and "changes" that I used with "uniformities" before. The latter works, so I'll try that. Finally, "predictable" must modify "changes" rather than "age" or "age level", since I can imagine charts somehow predicting changes more easily than their predicting ages."

Making this justification, as complicated as it seems, is relatively easy compared to the steps that were required to find it among all the other possibilities - formulating alternatives, making negative evaluations of some of them and differential choices among others, and finally settling on a chosen interpretation or several likely ones. Clearly some readers may boggle in the face of such a passage and give up. Others may have the processing capacity to carry through the kind of analysis outlined here, either entirely or partially below the level of introspection. Still others may adopt some control strategy that does not consider all of the alternatives. Some of them may do so with strategies that are still likely to obtain the correct interpretation most of the time (assuming that there is one), while others may adopt erroneous strategies that doom them to misunderstanding. Even these latter may find interpretations that are intellectually satisfying to themselves, causing them to assert that they understood the passage, although what they have understood may be almost totally unrelated to what the passage says.

To close our discussion of the interpretation of this second sentence, let me point out that a lot of the reasoning that was used for selecting an intended interpretation for "predictable age level changes" depends on the fact that the intended interpretation must fit the context "group growth charts show ..." in a passage whose topic is known to be growth (of something). Effectively, many different possible interpretations had to be hypothesized and tried in this context for a possible fit. It appears that below the level of conscious awareness, a great deal more hypothesis enumeration and evaluation is going on than one would first suspect. Current computer models tend to be based on the assumption that the amount of such hypothesis formation can be controlled by having the right dominating "frame" or "script", since the cost of considering many alternative hypotheses on a serial computer is prohibitive. I suspect, however, that this is one of the differences between the serial computer and the human brain that is significant and that in this respect the characteristics of computers as models of human processing is misleading. Much of AI work will continue to be focussed in this direction because of the desirability of getting computers to do such tasks, but I suspect that valid models of human performance will include much more parallel evaluation of alternative hypotheses.

10. Conclusion

The above discussion does not go into detail at the level of recognition of individual words and letters, but one can model them with similar processes, so that the overall reading process is a cascade of levels, each of which is making only tentative decisions. Each level will be formulating many alternative hypotheses that are to be partially selected by virtue of the degree to which they are compatible with hypotheses at other levels. For example, the syntactic component in the BBN speech understanding system makes many alternative hypotheses about possible syntactic paths through each of the theories that it is given by the control component to consider. Rumelhart [1977] gives a sketch of such a multi-level model based on the hypothesis structure of the Hearsay II speech understanding system [Lesser et al., 1975].

I have tried to give at a fairly concrete level some picture of the hypothesis formation and evaluation processes that must go on during reading although we are not normally aware of them. Many of these processes include inferences involving the kinds of metaknowledge discussed by A. Brown [1978], although the presentation here puts those processes in a somewhat different light. Here, I would stress that although we have some knowledge about what we know and how we know it, this knowledge is based on introspective observation in much the same way that our knowledge of any aspect of the world is based on observation. We do not in general have any privileged access to some "internal truth" in this respect. We do have an awareness of certain internal mental events that are perceivable in much the same way that our external sense organs perceive the world, but they do not give us a complete awareness of our internal mental processing, and we have to learn the significance of what awareness they do give us. This is presumably why the metacognitive abilities come rather late in the stage of mental development.

Experiments repeatedly show that what appears to be memory is in fact reconstruction and that many "memories" about which we are absolutely certain turn out to be mistaken. Thus, our metaknowledge of what we know and how we know it is only as good as the model we have built up based on our observations of our own performance and cannot generally be relied on as absolute truth. Conversely, a correct understanding at the meta level of how we should go about some mental process does not automatically translate into an ability to carry out that process. The attempt to consciously follow a set of instructions is not the same as fully incorporating those instructions into one's internal procedures. Examples of this phenomenon abound in such processes as learning to drive, learning to sail, learning to play chess, to solve mathematical problems, etc. The process whereby repeated attempts to follow such instructions eventually "compiles" an internal procedure for doing the task and the means whereby conscious resolution to "do it different next time" actually modifies such procedures are almost totally mysterious. It is the attempt at modeling such (non-introspectable) procedures by computer programs and abstract automata that I believe holds the key to understanding them, and it is this understanding which is the key to effective educational strategies.

I will not pretend that the results of Artificial Intelligence and Natural Language Processing research to date can give a complete account of the processes outlined above. However, they do provide a very rich inventory of analogies out of which one can construct hypothetical "brain computer architectures" and reasoning strategies that might model the information processing operations that go on in reading. The major contribution which the AI approach has to offer, I think, is that it reveals and makes concrete information processing steps that one might not otherwise have suspected. It can thus serve a very valuable function in the investigation of the reading process. On the other hand, one has to be careful in extrapolating results from computer models to human processes, since certain characteristics of any computer implementation will be determined by the nature of the computer on which they are implemented and may not be true of the "computer" in our heads.

In certain theoretical senses, investigations of abstract automata, such as Turing machines and abstract neural networks can tell us what kinds of functions various subparts of the brain might perform and what their limitations might be. However, we do not yet have a complete enough account of human intelligence in such terms to derive practical results. Humans are presumably heir to the same limitations that Turing machines are known to have in that they cannot possess algorithms to solve formally unsolvable problems, but beyond that we cannot begin to derive such useful predictions as how much new information per minute can humans learn, how many facts can they store and remember, or any of the myriad practical questions that one would like to know to design effective educational pedagogies.

What AI can do is serve a role very much like that which theoretical physics or chemistry serve for their respective fields. It can suggest models that have theoretical characteristics that fit the known data and predict unknown data. As such models converge and begin to be supported by empirical study, they can be put to a wide range of practical uses, such as designing pedagogical strategies and training material. However, the path between where we are and the ability to make such predictions will require a great deal of work.

Computer models up to and including the sentence level are now relatively well articulated and can be used quite well for analogies with human processing. Above the sentence level, however, current capabilities of computer systems are limited, and recognition of the intended interpretation of stories has a number of characteristics that make it fundamentally more difficult than individual sentence parsing. Most current attempts at this level rely on preselected scripts that constrain the possible interpretations of the sentences that they will encounter to a microscopic fraction of what could otherwise occur. Since human beings have encyclopedic amounts of knowledge that are drawn on and used in understanding what they read, something much more than the current script-based theories will be required to deal with general human behavior. However, increasing interest in this problem by linguists, computational linguists, philosophers, psychologists, and researchers in artificial intelligence give promise for the development of increasingly more adequate models of the overall reading process.

References

Becker, J.D. An Information-Processing Model of Intermediate-Level Cognition. BBN Report No. 2335, Bolt Beranek and Newman, Inc., Cambridge, Ma. (Ph.D. Thesis), 1972.

Brown, A.L. Metacognitive Development and Reading. In R. Spiro, B. Bruce and W. Brewer (eds.), Theoretical Issues in Reading Comprehension. Hillsdale, N.J.: Lawrence Erlbaum Assoc., 1978 (in press).

Bruce, B.C. Plans and Social Actions. In R. Spiro, B. Bruce and W. Brewer (eds.) Theoretical Issues in Reading Comprehension. Hillsdale, N.J.:Lawrence Erlbaum Assoc., 1978 (in press).

Collins, A.M. and Quillian, M.R. Tripping Down the Garden Path. BBN Report No. 2008, Bolt Beranek and Newman Inc., Cambridge, Ma. 1971.

Klatt, D.H. and Stevens, K.N. Strategies for Recognition of Spoken Sentences from Visual Examination of Spectrograms. BBN Report No. 2154, Bolt Beranek and Newman Inc., Cambridge, Ma., 1971.

Kuno, S. and Oettinger, A.G. Multiple-Path Syntactic Analyzer". Information Processing 1962, North-Holland Publishing Co., Amsterdam, 1963.

Lesser, V.R., Fennel, R.D., Erman, L.D. and Reddy, D.R. Organization of Hearsay II Speech Understanding System. IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-23, No. 1, February, 1975.

Nash-Webber, B.L. The Role of Semantics in Automatic Speech Understanding. In Representation and Understanding: Studies in Cognitive Science, D. Bobrow and A. Collins (eds.), New York: Academic Press, 1975.

Riverside Research Institute. The Reading Competency Test. The University of the State of New York, State Education Department, Albany, New York, Form B-1, 1974.

Rumelhart, D.E. Toward an Interactive Model of Reading. In S. Dornic (ed.), Attention & Performance VI. London: Academic Press, 1977.

Wanner, E. Do We Understand Sentences from the Outside-In or from the Inside-out? Daedalus, pp. 163-183, Summer, 1973.

Winograd, T. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. Project MAC Report, TR-84, MIT, 1971.

Wolf, J.J. and Woods, W.A. The HWIM Speech Understanding System. International Conference on Acoustics, Speech and Signal Processing (IEEE Catalog No. 77CH1197-3 ASSP), Hartford, Conn., May 9-11, 1977, 784-787.

Woods, W.A. Transition Network Grammars for Natural Language Analysis. Communications of the ACM, Vol. 13, No. 10, pp. 591-602, 1970.

Woods, W.A. Progress in Natural Language Understanding - An Application to Lunar Geology. AFIPS Proceedings, National Computer Conference and Exposition, 1973a.

Woods, W.A. Meaning and Machines. In Computational and Mathematical Linguistics, A. Zampolli (ed.), Proc. of the International Conference on Computational Linguistics, Pisa, Italy, 1973b. Florence, Italy: Leo S. Olschki, 1977.

Woods, W.A. Syntax, Semantics, and Speech. In D.R. Reddy (ed.), Speech Recognition: Invited Papers of the IEEE Symposium, Academic Press, pp. 345-400, 1975a.

Woods, W.A. What's in a Link: Foundations for Semantic Networks. In D. Bobrow & A. Collins (eds.)., Representation & Understanding: Studies in Cognitive Science. New York: Academic Press, 1975b.

Woods, W.A. and Makhoul, J.I. Mechanical Inference Problems in Continuous Speech Understanding. Artificial Intelligence, Vol. 5, No. 1, 73-91, 1974.

Woods, W.A., Bates, M., Brown, G., Bruce, B., Cook, C., Klovstad, J., Makhoul, J., Nash-Webber, B., Schwartz, R., Wolf, J., Zue, V. Speech Understanding Systems, Final Report, November 1974 - October 1976, BBN Report No. 3438, Bolt Beranek and Newman Inc., Cambridge, MA., 1976.

CENTER FOR THE STUDY OF READING

TECHNICAL REPORTS

No. 1: Halff, H. M. Graphical Evaluation of Hierarchical Clustering Schemes, October 1975.

No. 2: Spiro, R. J. Inferential Reconstruction in Memory for Connected Discourse, October 1975.

No. 3: Goetz, E. T. Sentences in Lists and in Connected Discourse, November 1975.

No. 4: Alessi, S. M., Anderson, T. H., & Biddle, W. B. Hardware and Software Considerations in Computer Based Course Management, November 1975.

No. 5: Schallert, D. L. Improving Memory for Prose: The Relationship Between Depth of Processing and Context, November 1975.

No. 6: Anderson, R. C., Goetz, E. T., Pichert, J. W., & Halff, H. M. Two Faces of the Conceptual Peg Hypothesis, January 1976.

No. 7: Ortony, A. Names, Descriptions, and Pragmatics, February 1976.

No. 8: Mason, J. M. Questioning the Notion of Independent Processing Stages in Reading, February 1976.

No. 9: Siegel, M. A. Teacher Behaviors and Curriculum Packages: Implications for Research and Teacher Education, April 1976.

No. 10: Anderson, R. C., Pichert, J. W., Goetz, E. T., Schallert, D. L., Stevens, K. V., & Trollip, S. R. Instantiation of General Terms, March 1976.

No. 11: Armbruster, B. B. Learning Principles from Prose: A Cognitive Approach Based on Schema Theory, July 1976.

No. 12: Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. Frameworks for Comprehending Discourse, July 1976.

No. 13: Rubin, A. D., Bruce, B. C., & Brown, J. S. A Process-oriented Language for Describing Aspects of Reading Comprehension, November 1976.

No. 14: Pichert, J. W., & Anderson, R. C. Taking Different Perspectives on a Story, November 1976.

No. 15: Schwartz, R. M. Strategic Processes in Beginning Reading, November 1976.

No. 16: Jenkins, J. R., & Pany, D. Curriculum Biases in Reading Achievement Tests, November 1976.

No. 17: Asher, S. R., Hymel, S., & Wigfield, A. Children's Comprehension of High- and Low-Interest Material and a Comparison of Two Cloze Scoring Methods, November 1976.

No. 18: Brown, A. L., Smiley, S. S., Day, J. D., Townsend, M. A. R., & Lawton, S. C. Intrusion of a Thematic Idea in Children's Comprehension and Retention of Stories, December 1976.

No. 19: Kleiman, G. M. The Prelinguistic Cognitive Basis of Children's Communicative Intentions, February 1977.

No. 20: Kleiman, G. M. The Effect of Previous Context on Reading Individual Words, February 1977.

No. 21: Kane, J. H., & Anderson, R. C. Depth of Processing and Interference Effects in the Learning and Remembering of Sentences, February 1977.

No. 22: Brown, A. L., & Campione, J. C. Memory Strategies in Learning: Training Children to Study Strategically, March 1977.

No. 23: Smiley, S. S., Oakley, D. D., Worthen, D., Campione, J. C., & Brown, A. L. Recall of Thematically Relevant Material by Adolescent Good and Poor Readers as a Function of Written Versus Oral Presentation, March 1977.

No. 24: Anderson, R. C., Spiro, R. J., & Anderson, M. Schemata as Scaffolding for the Representation of Information in Connected Discourse, March 1977.

No. 25: Pany, D., & Jenkins, J. R. Learning Word Meanings: A Comparison of Instructional Procedures and Effects on Measures of Reading Comprehension with Learning Disabled Students, March 1977.

No. 26: Armbruster, B. B., Stevens, R. J., & Rosenshine, B. Analyzing Content Coverage and Emphasis: A Study of Three Curricula and Two Tests, March 1977.

No. 27: Ortony, A., Reynolds, R. E., & Arter, J. A. Metaphor: Theoretical and Empirical Research, March 1977.

No. 28: Ortony, A. Remembering and Understanding Jabberwocky and Small-Talk, March 1977.

No. 29: Schallert, D. L., Kleiman, G. M., & Rubin, A. D. Analysis of Differences Between Oral and Written Language, April 1977.

No. 30: Goetz, E. T., & Osborn, J. Procedures for Sampling Texts and Tasks in Kindergarten through Eighth Grade, April 1977.

No. 31: Nash-Webber, B. Anaphora: A Cross-Disciplinary Survey, April 1977.

No. 32: Adams, M. J., & Collins, A. A Schema-Theoretic View of Reading Comprehension, April 1977.

No. 33: Huggins, A. W. F. Syntactic Aspects of Reading Comprehension, April 1977.

No. 34: Bruce, B. C. Plans and Social Actions, April 1977.

No. 35: Rubin, A. D. Comprehension Processes in Oral and Written Language, April 1977.

No. 36: Nash-Webber, B., & Reiter, R. Anaphora and Logical Form: On Formal Meaning Representations for Natural Language, April 1977.

No. 37: Adams, M. J. Failures to Comprehend and Levels of Processing in Reading, April 1977.

No. 38: Woods, W. A. Multiple Theory Formation in High-Level Perception, April 1977.

No. 39: Nickerson, R. S., & Adams, M. J. Uses of Context in Speech Understanding and Reading, April 1977.

No. 40: Brown, J. S., & Collins, A. Model-Based Versus Text-Based Reasoning, April 1977.

No. 41: Anderson, R. C., & Pichert, J. W. Recall of Previously Unrecallable Information Following a Shift in Perspective, April 1977.

No. 42: Mason, J., Osborn, J., & Rosenshine, B. A Consideration of Skill Hierarchy Approaches to the Teaching of Reading, April 1977.

No. 43: Collins, A., Brown, A. L., Morgan, J. L., & Brewer, W. F. The Analysis of Reading Tasks and Texts, April 1977.

No. 44: McClure, E. Aspects of Code-Switching in the Discourse of Bilingual Mexican-American Children, April 1977.

No. 45: Schwartz, R. M. Relation of Context Utilization and Orthographic Automaticity in Word Identification, May 1977.

No. 46: Anderson, R. C., Stevens, K. C., Shifrin, Z., & Osborn, J. Instantiation of Word Meanings in Children, May 1977.

No. 47: Brown, A. L. Knowing When, Where, and How to Remember: A Problem of Metacognition, June 1977.

No. 48: Brown, A. L., & DeLoache, J. S. Skills, Plans, and Self-Regulation, July 1977.

No. 49: Goetz, E. T. Inferences in the Comprehension of and Memory for Text, July 1977.

No. 50: Anderson, R. C. Schema-Directed Processes in Language Comprehension, July 1977.

No. 51: Brown, A. L. Theories of Memory and the Problems of Development: Activity, Growth, and Knowledge, July 1977.

No. 52: Morgan, J. L. Two Types of Convention in Indirect Speech Acts, July 1977.

No. 53: Brown, A. L., Smiley, S. S., & Lawton, S. C. The Effects of Experience on the Selection of Suitable Retrieval Cues for Studying from Prose Passages, July 1977.

No. 54: Fleisher, L. S., & Jenkins, J. R. Effects of Contextualized and Decontextualized Practice Conditions on Word Recognition, July 1977.

No. 55: Jenkins, J. R., & Larson, K. Evaluating Error Correction Procedures for Oral Reading, August 1977.

No. 56: Anderson, T. H., Standiford, S. N., & Alessi, S. M. Computer Assisted Problem Solving in an Introductory Statistics Course, August 1977.

No. 57: Barnitz, J. Interrelationship of Orthography and Phonological Structure in Learning to Read, August 1977.

No. 58: Mason, J. M. The Role of Strategy in Reading in the Mentally Retarded, September 1977.

No. 59: Mason, J. M. Reading Readiness: A Definition and Skills Hierarchy from Preschoolers' Developing Conceptions of Print, September 1977.

No. 60: Spiro, R. J., & Esposito, J. J. Superficial Processing of Inferences Presented in Text, October 1977.

No. 61: Spiro, R. J., & Smith, D. Distinguishing Sub-Types of Poor Comprehenders: Overreliance on Conceptual vs. Data-Driven Processes, October 1977.

No. 62: Spiro, R. J., & Rittenhouse, R. K. Attribute Clarity in Conservation Instructions and Horizontal Décalage in Normal Hearing and Deaf Children, October 1977.

No. 63: Spiro, R. J., & Anderson, M. C. Theme and Subordination in Sentence Recall, October 1977.

No. 64: Spiro, R. J., & Martin, J. E. Contextual Factors in the Recall of Alternative Surface Structures, October 1977.

No. 65: Brewer, W. F. Memory for the Pragmatic Implications of Sentences, October 1977.